

UNITED STATES PATENT APPLICATION
FOR

ISOCHRONOUS QUEUE AND BUFFER MANAGEMENT

INVENTORS:

PAULINE YEUNG

PREPARED BY:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN, LLP
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1026

(408) 720-8598

"Express Mail" mailing label number: EL431887034US
Date of Deposit: April 7, 2000

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, D. C. 20231

Carla Zavala

(Typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

(Date signed)

004509.P010

ISOCHRONOUS QUEUE AND BUFFER MANAGEMENT

FIELD OF THE INVENTION

The present invention relates generally to switches. More specifically, the present invention relates to isochronous queue and buffer management in switches.

BACKGROUND OF THE INVENTION

The IEEE Standard for a High Performance Serial Bus, IEEE Std. 1394-1995 published August 30, 1996 (1394-1995 Standard) and its progeny provide a high speed serial protocol which permits implementation of high speed data transfers. The existing progeny includes P1394a Draft Standard for a High Performance Serial Bus (1394a Standard) and P1394b Draft Standard for a High Performance Serial Bus (1394b Standard). Generically, systems implementing 1394-1995, 1394a, 1394b or subsequent revisions and modifications thereof are referred to herein as 1394 systems.

The IEEE 1394 standard is an international standard for implementing a high-speed serial bus architecture, which supports both asynchronous and isochronous format data transfers. The IEEE 1394 standard defines a bus as a non-cyclic interconnect. Within a non-cyclic interconnect, devices may not be connected together so as to create loops.

In networks, switches filter and forward packets between local area network segments. In packet switching, packets are individually routed between nodes with no previously established communication path. An algorithm is used to route packets to their destination through the most expedient route. The destination computer reassembles the packets in their appropriate order. Packet switching optimizes the use of bandwidth

available in a network and minimizes the latency (the time it takes for a packet to cross a network connection, from sender to receiver).

In a 1394 network with multiple 1394 buses and 1394 switches, all the 1394 buses should be synchronous. But due to cycle skewing, the cycle start packets are not all

5 generated at the same time in different 1394 buses. Cycle skewing occurs when a large asynchronous packet is sent over a bus and the large packet is late, which may delay the start of the next cycle.

In a switch, there may be packets arriving from different ingress ports routed to one egress port. Because of cycle skewing, a packet from one cycle may arrive in the
10 egress port after a packet from a subsequent cycle.

Packets being switched may also be transmitted out of order from the switch because a first packet arriving before a second packet at a switch may not be completely received before the second packet is completely received. Thus, the second packet would
15 be sent out before the first packet because the second packet was completely received before the first packet.

SUMMARY OF THE INVENTION

A method of processing packets in a switch is described. A first queue is selected from at least three queues based on the cycle number (C) of a cycle and flushed at the start of cycle C. At least one isochronous packet is received over a bus during the cycle.

- 5 The packet is placed in a second queue based on the cycle number.

004040-5587550

BRIEF DESCRIPTION OF THE DRAWINGS

Features and advantages of the prevention invention will be apparent to one skilled in the art in light of the following detailed description in which:

Figure 1 is a block diagram of one embodiment of a switch in a communications
5 network;

Figure 2 is a block diagram of one embodiment of a switch;

Figure 3 is a block diagram of packet processing in one embodiment of a switch;

Figure 4 is a block diagram of buffer management in one embodiment of a
switch;

Figure 5 is a block diagram of delay adjustment in one embodiment of a switch;
10 and

Figure 6 is a flow diagram for packet processing in one embodiment of a switch.

004040-59874550

DETAILED DESCRIPTION

ing
A method and system for a method and system for isochronous queue and buffer management are described.

In the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However, it will be apparent to one skilled in the art that the present invention may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

002010-5934560
10 **Figure 1** is a block diagram of one embodiment of the switch in a communications network. Switch 20 is connected to devices 1-5, a wide area network 7 and a local area network 6. Devices 1-5 may include audio, video and/or audio/video devices including storage systems and telecommunications. The wide area network may include the internet or proprietary network or a television communications network.

15 **Figure 2** is a block diagram of one embodiment of a switch. **Figure 2** shows a switch 220 including a processor 228 and a buffer 229. The processor 228 directs operations within the switch 220 and the buffer 229 stores switched packet streams to be transmitted as described below. Switch 220, according to one embodiment, is configured to switch packets on a IEEE 1394 Standard Serial Bus.

20 **Figure 3** is a block diagram of packet processing in one embodiment of a switch. Figure 3 illustrates a switch 320 having ingress ports 321-324 and egress ports 331-334. The number of ingress ports and egress ports in the switch may vary depending on the application and how many devices or buses are served by the switch.

Each ingress port 321-324 receives a stream of packets. Ingress port 321 receives a stream of packets including packets 311a-b, 312a-b, and 313a-b. Ingress port 322 receives a stream of packets including packets 311c-d, 312c-d, and 313c-d. Ingress port 323 receives a stream of packets including packets 311e-f, 312e-f, and 313e-f. Ingress port 324 receives a stream of packets including packets 311g-h, 312g-h, and 313g-h.

Each ingress port 321-324 may be associated with an IEEE Standard 1394 Bus (not shown) or other connections or channels including ethernet, asynchronous transfer mode (atm), T-1 or T-3 carrier, OC-X or any other suitable connection. Packets in packet streams 311-313 may be isochronous packets, according to one embodiment, or any other type of packet that is suitable.

Figure 3 shows packet streams 311-313 arriving at ingress ports 321-324 as follows: packets 311a-h arrive at ingress ports 321-324 during cycle (N) 351, packets 312a-h arrive at ingress ports 321-324 during cycle (N + 1) 352, and packets 313a-h arrive at ingress ports 321-324 during cycle (N + 2) 353.

Figure 3 shows packet stream 311 leaving switch 320 through egress ports 331-334 at cycle (N + 2) 353. Packets 311 arriving at switch 320 at cycle N may be switched during cycles (N), (N + 1) and, possibly (N + 2). The packets 311 are sent out during cycle (N + 2), two cycles after they arrive. As shown in egress cycles 351 and 352, packets 309, which arrived during cycle (N - 2) (not shown) are sent out at cycle (N) 351 and packets 310, which arrived during cycle (N - 1) (not shown) are sent out at cycle (N + 1) 352.

The minimum delay of a 1394 isochronous packet in a 1394 switch is two 1394 cycles. Thus, the 1394 switch 320 uses a buffer management system, as discussed below, to assure that packets arriving at a cycle (N) 351 are sent out in a cycle (N + 2) 353. *flow*

Figure 4 is a block diagram of a buffer management system in one embodiment of a switch. In the embodiment shown, buffer management system 450 includes four queues Q0-Q3 451-454 each having a Used pointer 451-454a and a Free pointer 451-454b. Although the buffer management system 450 shown in **figure 4** includes four queues, three queues may be used instead. Each egress port 331-334 has a buffer management system 450 including four queues Q0-Q3 451-454.

In the embodiment shown, packet streams 411, 412, and 413 have arrived at a switch 320. Packet streams 411-413 include packets 411a-411d, 412a-412d and 413a-413d, respectively. The queue numbers correspond to the cycle of the switch 320 in which the packet streams 411-413 will be sent out. As shown by block 435, a packet stream P0(C0) 413, where (C0) represents a time stamp of cycle C0, arrives at cycle C0 (and will be sent out at cycle C2) *C2*

As egress packet queues 451-454 are filled up, the free pointer values 451-454b are set at free = n, where n represents the point at which packets may be added to queues 451-454. Thus, as shown with reference to Q2 453, free pointer 453b points to the free space after packet 413c, where packet 413d is to be received. *Key*

Also, as the packet queues 451-454 are filled up, the used pointer value is set to represent the next packet to be transmitted, as shown by pointers 451-454a. As shown with reference to Q2 453, used pointer 453a is set to 0 and points to space from which the next packet is to be transmitted from the queue Q2 453, through an egress port 331-334 *Key*

of switch 320. With reference to Q0 451, which is in the process of transmitting packets, the used pointer value 451a equals m, where m represents the space from which the next packet will be sent.

When an packet queue 451-454 is flushed or cleared, a used pointer value 451-454b is set to 0 to show that the space from which the next packet is to be transmitted. As shown with reference to Q3 454, used pointer 454b points to the space from which the next packet is to be transmitted from queue Q3 454, through an egress port 331-334 of switch 320. Also, when an packet queue 451-454 is flushed or cleared, a free pointer value is set to 0 to show that the queue is empty. As shown with reference to Q3 454, to show that the queue 454 is free, 454a is set to 0 to point to the place in queue Q3 454 where the next packet may be placed.

At any given cycle C of the switch 320, there are isochronous packets in the switch 320 which arrived in cycle C, C - 1 557, and C - 2 556. To guarantee the packets are sent in the proper cycle, at least three packet queues are needed for every egress port.

Figure 5 is a block diagram of delay adjustment in one embodiment of a switch. A delay adjustment may become necessary where cycle skewing, as described above, occurs. A switch 320 includes an inbound cycle 525, a transition cycle 526 and an outbound cycle 527. Packet streams 511a-d, 512a-d, 513a-d, and 514a-d are received at switch 320 during cycles C - 3 555, C - 2 556, C - 1 557, and C 558 of the inbound cycle of the switch.

In the embodiment shown, input packets 511 comes in at inbound cycle C - 3 555 and input packets 512 starts to come in at inbound cycle C - 2 556 of inbound cycle 525. However input packet stream 512 includes packet 512d which comes in at the end of

cycle C - 2 556 and almost in cycle C - 1 557. Thus, cycle C - 2 556 is expanded from 556a to accommodate the late packet and cycle C - 1 557 is shortened.

Expand/contract cycle time to Account for different lengths of packets

In transition cycle 526, packets 512a and 512b of input packets 512 are switched during cycle C - 2 556 and placed in the appropriate egress packet queue 451-454, while

5 packets 512c and 512d are switched at cycle C - 1 557 and placed in the appropriate packet queue 451-454. Because packets 512 arrived during inbound cycle C - 2 556, it will be buffered to go out at cycle C ^q even though a few of the packets were late. Thus, input packets 512 is sent out at cycle C 558 of outbound cycle 527.

Key

Packets 511, received at cycle C - 3 555 is switched during cycle C - 2 556 of the transition cycle, and sent out at cycle C - 1 557. However, packets 511 need not be switched at cycle C - 2 556. Packets 513, which is received during cycle C - 1 557 is switched during cycles C - 1 557 and C 558, and will be sent out during cycle C + 1, not shown. Packet 513a is switched during cycle C - 1 557 and packets 513b-d are switched during cycle C 558. Also, as shown, packets 514a-b are switched at cycle C 558, the cycle during which packets 514 was input.

Figure 6 is a flow diagram for packet processing in one embodiment of a switch.

At processing block 671, a first queue is selected based on the cycle number of the egress cycle of the switch 220. At processing block 672, the first queue is flushed at the start of

the cycle. At processing block 673, an isochronous packet is received over a bus. At

20 processing block 674, the packet is placed in a packet queue based on the cycle number of

the cycle. The multiple packet queues Q0-Q3 451-454 are used in each egress port of the 1394 switch to resolve the order of packets that will depart in different cycles according to the egress 1394 cycle time.

Push into the queues like queues Q0, Q1 & send the data out over the bus based on the cycle # of the bus

distorted due to skewing

ultimate key

Thus, if n queues are used for each port, where $n \geq 3$, packets that arrived in cycle

multi C would go to queue number $(C + 2) \% n$ (where $\%$ stands for remainder). In cycle C, the

send egress port sends packets from queue $(C \% n)$. Thus, if cycle number C is 5 and the

number of queues equals 4, the egress port will send packets from queue number Q1,

5 since the remainder of $(5 / 4)$ is one.

*in cycle # 5: $(5+2) \% 4 = 1$
 $7 \% 4 = 3$ one*

When a packet stream 411 arrives at the egress port, as represented in processing block 673, a packet buffer 451-454 is allocated from the packet buffer pool to hold the complete packet. When this packet departs, there is no need to free the memory associated with this packet buffer. Rather at the beginning of each cycle C, all memory in

10 the packet buffer pool associated with packet queue number $(C - 1) \% n$ is reclaimed.

Thus if the cycle number is 5 and the number of queues equals 4, Q0 451 would be flushed at the beginning of cycle five while Q1 is being filled up. Thus, the queue being filled, the queue being flushed and the queue having packets transmitted are all based on the cycle number of the 1394 switch.

15 In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes can be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

Done to multiple packets at once!